



Requested Patent: JP2001209565A

Title: ASYNCHRONOUS REMOTE COPY ;

Abstracted Patent: EP1120711 ;

Publication Date: 2001-08-01 ;

Inventor(s):

NAKANO TOSHIO (JP); TABUCHI HIDEO (JP); NOZAWA MASAFUMI (JP); SHIMADA AKINOBU (JP) ;

Applicant(s): HITACHI LTD (JP) ;

Application Number: EP20000105491 20000315 ;

Priority Number(s): JP20000024687 20000128 ;

IPC Classification: G06F11/14 ;

Equivalents: ;

ABSTRACT:

To provide a disk subsystem that assures the sequence and the coherence of data update regarding two or more disk subsystems, and to provide an asynchronous type remote copy function, a main center (9) that is a computer system having the configuration of slave subsystems (3-2 to 3-n) connected with a master disk subsystem (3-1) secures the coherence between data of the main center (9) and the remote center (10) at the temporary suspension by repeating temporary suspension and release of the temporary suspension of the remote copy by a master subsystem (3-1) at predetermined opportunities, and by repeating temporary suspension and release of the temporary suspension of the remote copy by slave subsystems (3-2 to 3-n) interlocking with it.

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2001-209565

(P2001-209565A)

(43)公開日 平成13年8月3日(2001.8.3)

(51)Int.Cl'	識別記号	F I	テ-マコード*(参考)
G 0 6 F 12/00	5 3 3	G 0 6 F 12/00	5 3 3 J 5 B 0 6 5
	5 3 1		5 3 1 D 5 B 0 8 2
3/06	3 0 4	3/06	3 0 4 F

審査請求 未請求 請求項の数17 OL (全 11 頁)

(21)出願番号	特願2000-24687(P2000-24687)	(71)出願人	000005108 株式会社日立製作所 東京都千代田区神田駿河台四丁目6番地
(22)出願日	平成12年1月28日(2000.1.28)	(72)発明者	田渕 英夫 神奈川県小田原市国府津2880番地 株式会社日立製作所ストレージシステム事業部内
		(72)発明者	島田 朗伸 神奈川県小田原市国府津2880番地 株式会社日立製作所ストレージシステム事業部内
		(74)代理人	100075096 弁理士 作田 康夫

最終頁に続く

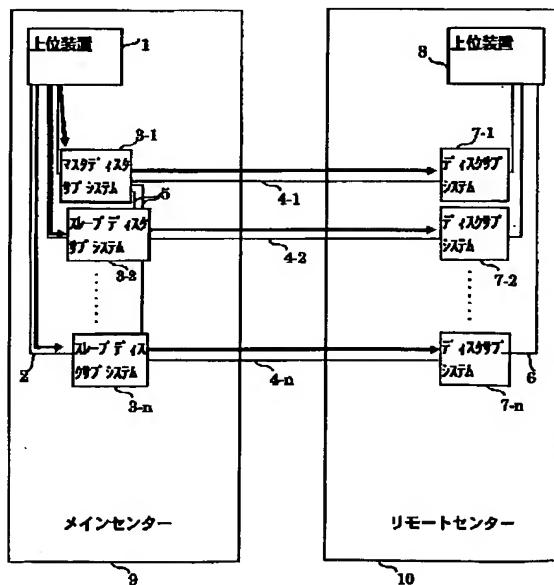
(54)【発明の名称】 サブシステム及びこれらの統合システム

(57)【要約】

【課題】複数のディスクサブシステムにわたってデータ更新の順序性、整合性を保証できる、非同期型のリモートコピー機能を有するディスクサブシステムを提供する。

【解決手段】メインセンター9では、一台のマスタサブシステム3-1にスレーブサブシステム3-2～3-nが接続されているコンピュータシステム構成であって、予め設定された契機でマスタサブシステム3-1がリモートコピーの一時停止停止と一時停止解除を繰り返し、これに連動してスレーブサブシステム3-2～3-nがリモートコピーの一時停止停止と一時停止解除を繰り返すことで一次停止におけるメインセンター9とリモートセンター10のデータの整合性を取る。

図1



【特許請求の範囲】

【請求項1】上位装置とデータの授受を行う制御手段と、前記データの格納を行う記憶手段と、前記データを離れた場所に在るディスクサブシステムに予め定めた間隔でデータ転送/転送中断を行う手段とを有するマスタディスクサブシステムと、

それぞれが、前記上位装置とデータの授受を行う制御手段と、前記データの格納を行う記憶手段と、前記マスタディスクサブシステムに接続される伝送手段と、この伝送手段を使用して前記マスタディスクサブシステムがデータ転送状態か転送中断状態かを問い合わせる問い合わせ手段と、この問い合わせの結果マスタディスクサブシステムがデータ転送状態である時に前記データを離れた場所に在るディスクサブシステムに転送する手段とを有する複数のスレーブディスクサブシステムと、前記マスタ及びスレーブディスクサブシステムとは離れた場所に在り、前記各サブシステムから転送されるデータを受け取る制御手段と、このデータの格納を行う記憶手段とを有する複数のリモートディスクサブシステムとを有し、

前記マスタ及びスレーブディスクサブシステムは前記リモートサブシステムにデータを間欠的に転送する事で、転送中断時における前記マスタ及びスレーブサブシステム内のデータと前記リモートディスクサブシステム内のデータとの一貫性を確保する統合システム。

【請求項2】スレーブサブシステムのリモートコピー方法であって、

上位装置からデータを受取り、このデータを記憶手段に格納し、マスタサブシステムがデータをリモートセンターに転送する状態をマスタサブシステムに問い合わせ、前記マスタサブシステムが転送する状態の場合には前記データをリモートセンターに転送し、前記マスタサブシステムが転送しない状態の場合には前記データの格納位置に関する情報を保持し、前記マスタサブシステムが転送する状態となってからリモートセンターに転送するリモートコピー方法。

【請求項3】上位装置に対し情報の授受を行う制御部と、前記情報の格納を行う記憶部と、前記上位装置からの前記情報を他のサブシステムに設定した契機で一時停止しつつ転送する転送手段を備えるサブシステム。

【請求項4】前記サブシステムは、前記一時停止とこの一時停止の解除を周期的に行う請求項3に記載のサブシステム。

【請求項5】前記一時停止時には、転送すべき更新情報の格納位置情報を自サブシステム内に一時停止解除まで保持する請求項3に記載のサブシステム。

【請求項6】上位装置と情報の授受を行う制御部と、前記情報の格納を行う記憶部と、前記上位装置からの前記情報を他の場所に在るサブシステムに転送する転送部とを備える第一のサブシステムと、

上位装置と情報の授受を行う制御部と、前記情報の格納を行う記憶部と、前記上位装置からの前記情報を他の場所に在るサブシステムに転送する転送部と、前記マスタサブシステムと接続する伝送経路とを備える複数のサブシステムとからなる統合システム。

【請求項7】前記複数のサブシステムは、前記伝送経路を用いて前記第一のサブシステムの転送状態を入手する請求項6に記載の統合システム。

【請求項8】上位装置とデータの授受を行う制御部と前記データの格納を行う記憶部とをそれぞれに有し、前記上位装置からの前記データを他の場所に在るディスクサブシステムに転送するディスクサブシステムを複数備えた統合システムであって、前記上位装置に接続される複数のディスクサブシステムのうち第一のディスクサブシステムに対して、他のディスクサブシステムが前記他の場所に在るディスクサブシステムへの前記データの転送可否を問い合わせる統合システム。

【請求項9】前記転送の可否は、前記第一のディスクサブシステムが転送を一時停止しているか否かで判断する請求項8に記載の統合システム。

【請求項10】前記第一のディスクサブシステムは、前記一時停止とこの一時停止の解除を周期的に行う請求項9に記載の統合システム。

【請求項11】前記上位装置に接続された前記ディスクサブシステムと前記他の場所に在るディスクサブシステムとのデータ転送は、非同期型で行う請求項8に記載の統合システム。

【請求項12】上位装置に接続された第1のディスクサブシステム群と、この第1のディスクサブシステム群からデータの転送を受ける第2のディスクサブシステム群を有する統合システムであって、

前記第1のディスクサブシステム群のディスクサブシステムの中で1つは、前記第1のディスクサブシステム群が前記第2のディスクサブシステム群へデータ転送する際ににおける、転送の可否を管理するマスタサブシステムである統合システム。

【請求項13】前記転送の可否は、前記マスタサブシステムが前記第2のディスクサブシステム群へのデータ転送を一時停止しているか否かで判断する請求項12に記載の統合システム。

【請求項14】前記第1のディスクサブシステム群と前記第2のディスクサブシステム群との間は、SAN (Storage Area Network) を介して接続されている請求項12に記載の統合システム。

【請求項15】前記第1のディスクサブシステム群と前記第2のディスクサブシステム群との間のデータ転送は、非同期型で行う請求項12に記載の統合システム。

【請求項16】前記非同期型のデータ転送に際し、更新順序保証の要否を専用のディスクボリュームを指定する

ことによって明示する請求項15に記載の統合システム。

【請求項17】上位装置からデータを受取り、このデータを記憶手段に格納し、マスタサブシステムがデータをリモートセンタに転送する状態か否かをマスタサブシステムに問い合わせることでマスタサブシステムとの一貫性を確保する一貫性保全方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明はコンピュータシステムのデータを格納する外部記憶装置及びこれらの統合システムに関し、特に、遠隔地に存在する複数の外部記憶装置群（サブシステム）と、他の複数の外部記憶装置群とを相互に接続し、上位装置たるホストコンピュータを経由せずに、遠隔地に存在する外部記憶装置（サブシステム）との間で、データを二重化するリモートコピー技術に関する。ここで、サブシステムとは、上位装置に対し情報の授受を行う制御部と、情報の格納を行うディスク装置等を内蔵する記憶装置をいうものであり（記憶装置がディスク装置の場合はディスクサブシステムのように言う）、統合システムとは複数のサブシステムから構成されるシステムをいうものとする。

【0002】

【従来の技術】地震等の天災の際のデータのバックアップを考慮すれば、メインセンターとリモートセンターは100km～数100km程度、分離する必要がある。このため、メインセンターとリモートセンターにそれぞれ設置されているサブシステムの間で、データを二重化して保有する、いわゆる、リモートコピー機能を採用した外部記憶システムが、既にいくつか実用化されている。

【0003】リモートコピー機能は、同期型と非同期型の2種類に大別される。

【0004】同期型とはメインセンター内のホストコンピュータ（上位装置）からサブシステムに、データの更新（書き込み）指示が有った場合、その指示対象がリモートコピー機能の対象でもあるときは、そのリモートコピー機能の対象であるリモートセンターにおけるサブシステムに対して、指示された更新（書き込み）が終了してから、メインセンターの上位装置に更新処理の完了を報告する処理手順をいう。この場合、メインセンターとリモートセンターとの地理的距離に応じて、この間に介在するデータ伝送線路の能力の影響を受け、時間遅れ（伝送時間等）が発生する。

【0005】これに対し非同期型とは、メインセンター内の上位装置からサブシステムに、データの更新（書き込み）指示が有った場合、その指示対象がリモートコピー機能の対象であっても、メインセンター内のサブシステムの更新処理が終わり次第、上位装置に対し更新処理の完了を報告し、リモートセンターのサブシステムにお

けるデータの更新（反映）はメインセンターにおける処理とは非同期に実行する処理手順をいう。このためメインセンター内部で必要とされる処理時間でデータ更新が終了するので、リモートセンターへのデータの格納に起因する伝送時間等はかかるない。

【0006】非同期型は、リモートセンターのサブシステムの内容が、メインセンター側のそれに対し、常に一致しているわけではない。このため、メインセンターが災害等により機能を失った場合は、リモートセンター側にデータの反映が完了していないデータが消失することとなる。しかし、メインセンター側のサブシステムのアクセス性能を、リモートコピー機能を実施しない場合と同等レベルとすることができます。

【0007】かかる従来技術では、次の様な課題があった。

【0008】「一貫性保全について」リモートコピーを行う場合、メインセンターのサブシステムとリモートセンターのサブシステム間は独立した通信リンクで接続される。つまり、メインセンターの複数のサブシステムとリモートセンターの複数のサブシステムの間でリモートコピーを行う場合は、独立した通信リンクで接続されたサブシステムの組が複数存在する構成となる。

【0009】この構成では、リモートセンターへのコピーは各通信リンク毎に行われるため、サブシステムによりデータの更新された時刻が異なる。そのため、メインセンターが機能を失った場合に、どの時点までのデータがリモートセンターにコピーされていたのかが各サブシステム毎に異なるため、データ毎にタイムスタンプ等が付けられていない場合、メインセンターとリモートセンターの整合性を取ることが出来なかった。

【0010】このように、複数のサブシステムを有するメインセンターのバックアップをリモートセンターで行う場合、複数のサブシステム間でデータの一貫性を保全する（更新順序を保持する）という課題がある。尚、ここにおける「順序」とは、単なるデータの順番ではなく、リモート側の各サブシステムにおけるデータ更新の一貫性がある「状態」という意味で使用している。

【0011】非同期型リモートコピーでは、リモートセンターへの更新データの反映が、メインセンターでの実際の更新処理の発生時点より遅れて処理されることはやむを得ない。しかし更新の順序はメインセンターと一致していかなければならない。もしくは間欠的に見てメインセンターで発生した更新処理の順序とリモートセンターの更新処理の順序は整合性の取れた状態になければならない。

【0012】一般にデータベース等はデータベース本体と各種ログ情報、制御情報から構成されており、それぞれが関連性を持っている。データ更新の際はデータベース本体に加え、これらログ情報、制御情報をも更新し、システムの整合性が保たれている。したがって更新の順

序が崩れた場合、更新順序に関連するこれらの情報の整合性も崩れ、最悪の場合には、データベース全体の破壊につながる可能性がある。

【0013】「上位装置が介在することについて」メインセンター及びリモートセンターに複数のサブシステムが存在する一般的な環境で非同期型のリモートコピーを実現する場合には、上位装置がサブシステムへデータの更新を指示する場合、タイムスタンプなどの更新順序に関する情報をデータに付加し、これらの情報に基づいて副側のサブシステムの更新データ反映処理が実行されるのが一般的である。例えば、特開平6-290125号公報（米国特許第5446871号）に示されるリモートコピーの一様に、上位装置が介在してリモートコピー機能を実現している。

【0014】特開平6-290125号公報（米国特許第5446871号）に開示されている技術においては、メインセンター側の上位装置のオペレーティングシステムとディスクサブシステム、リモートセンター側の上位装置のデータマーバーソフトウェアとディスクサブシステムの連携により、更新順序情報の発行、送付、これに基づく更新データ反映処理を実現している。

【0015】

【発明が解決しようとする課題】従来技術により、メインセンター、リモートセンター間の更新順序性を保証しながら非同期型のリモートコピー機能が実現できる。しかし従来技術では、上位ソフトウェアとサブシステムの双方にこの機能実現の為の仕組みが必要であり、且つ、両者が連携しなければならない。そのため専用の新規ソフトウェアの導入が必要となり、ユーザは、ソフトウェアの導入、設定、検査、CPU負荷増加に伴うシステム設計の見直し等の作業が発生する。このためこの機能を導入するためにはかなりの期間を要し、費用が発生してしまう。

【0016】また、サブシステムの機能のみで非同期型のリモートコピー機能を実現する場合は、サブシステムの機能のみでデータの更新順序の整合性を保持する必要がある。更新順序の整合性が必要なデータが複数のサブシステムに分散されて格納されている場合、複数のサブシステム間で更新順序の整合性を保持するための手段がないという課題もあった。また、更新順序を保持することによるサブシステム内部の処理オーバヘッドの増加を防ぐことも課題の一つである。

【0017】本発明の目的は、上位装置に対して新規ソフトウェアの導入を必要とせず、サブシステムの機能で、データの整合性を保証でき、導入が容易かつメインセンターの性能低下が少ない、非同期型のリモートコピー機能を実現することである。

【0018】

【課題を解決するための手段】相互に遠隔地に存在するメインセンターとリモートセンターのサブシステム同士

を接続する。サブシステムが複数の場合は、メインセンターにおいて、サブシステムのうちの一台に（以下、マスタサブシステムと呼ぶ）、他のサブシステム（以下、スレーブサブシステムと呼ぶ）をすべて接続する。

【0019】メインセンターのマスタサブシステムは、リモートコピーの一時停止や一時停止の解除を行う。これらの動作契機はユーザがマスタサブシステムに予め設定しておく事により、リモートコピー実施中の任意の時点で実施する。メインセンターのサブシステムは上位装置より更新データを受け取ると、自サブシステムへのデータの格納を開始する。

【0020】そして、マスタサブシステムは自サブシステムがリモートコピーを一時停止した状態になっているか否かを確認し、一時停止状態になっていない場合は、当該データをリモートセンターへの送付対象とする。また、スレーブサブシステムは、ホストからの更新データを受け取った際にマスタサブシステムの状態を後述する制御ビットにて確認し、マスタサブシステムがリモートコピーを一時停止させる状態になっていないことを確認した場合のみ更新データをリモートセンターへの送付対象とする。

【0021】スレーブサブシステムがマスタサブシステムの状態を確認した際に、マスタサブシステムがリモートコピーを一時停止する状態になっていると、スレーブサブシステムはマスタサブシステムと同様にリモートコピーを一時停止する状態となる。一時停止状態では、マスタサブシステム、スレーブサブシステムを問わず、リモートセンターへのデータ送付は行わず、メインセンター側でのみ更新処理を行い、更新情報を各サブシステム内で管理・保存する。マスタサブシステムにおいて、リモートコピーの一時停止状態が解除された後は、マスタサブシステム、スレーブサブシステムともリモートセンターへのデータ送付を再開する。この場合、一時停止状態の間にメインセンタで更新されたデータも送付データの対象となる。

【0022】このようにして、メインセンターとリモートセンターの各サブシステムの各ボリュームの間は、非同期型のリモートコピーによりデータの二重化を行う。

【0023】

【発明の実施の形態】以下、図面を参照しながら本発明をコンピュータシステムに適用した場合の一例について説明する。

【0024】図1に、コンピュータシステムを装備した複数のデータセンターにおいて、任意の2つのセンター間で情報（データ）の二重化を行うために、本発明を適用したときの構成例を示す。

【0025】メインセンター9側の一台又は複数台のディスクサブシステム3と、リモートセンター10側の一台又は複数台のディスクサブシステム7は、上位装置（ホストコンピュータ）1、8を介さずに接続され、両

センター間でデータの二重化を行うリモートコピーシステムを実現している。上位装置を介さないディスクサブシステムの接続としては、例えばSAN (Storage Area Network) が挙げられる。図4にメインセンター9のディスクサブシステム3の構成例を示す。

【0026】図1のメインセンター9において、データ処理を行う中央処理装置(CPU)を持つ上位装置1は、伝送経路であるインタフェースケーブル2を介して、ディスクサブシステム3-1(マスタサブシステム)、3-2、……3-n(スレーブサブシステム)に接続されている。

【0027】ディスクサブシステム3-1、3-2、……3-nは、上位装置1からのデータ(情報を含む)授受を行うインターフェース制御部11と、上位装置1から参照又は更新されるデータデータを格納するバッファ12と、リモートコピーが一次停止中の更新データの格納位置に関する情報を格納するリモートコピー制御情報格納部16と、このデータを記録する記録媒体としての磁気ディスクドライブ13、これらのデータのやり取りを制御するマイクロプロセッサ14、これらの各要素を制御するディスクアレイサブシステム制御部17を備える。

【0028】また、マスタディスクサブシステム3-1はそれに加えてリモートコピーをどのような設定にて行うかをユーザが設定するサービスプロセッサパネル15を備える。そしてリモートコピー制御情報格納部16は、リモートコピーが一次停止中の更新データの格納位置に関する情報に加えサービスプロセッサパネル14により設定された制御情報より現在のリモートコピー状況を表す制御ビットを格納する。

【0029】マスタディスクサブシステム3-1はインターフェースケーブル4-1を介してリモートセンタ10のディスクサブシステム7-1と接続される。同様に、スレーブディスクサブシステム3-2はインターフェースケーブル4-2を介してリモートセンタのディスクサブシステム7-2と接続され、スレーブディスクサブシステム3-nはインターフェースケーブル4-nを介してリモートセンタのディスクサブシステム7-nと接続される構成をとる。

【0030】なお、インターフェースケーブル4-1、4-2、……4-nは、回線接続装置等を利用して一般的の通信回線と接続することも可能である。本例ではこの点も含めてインターフェースケーブル4-1～nとして記述する。

【0031】また、ディスクサブシステム3が複数台有る場合には、ディスクサブシステム3-1はメインセンター9内でリモートコピー対象のデータが格納されるディスクサブシステム3-1以外のディスクサブシステム3-2、……3-nとインターフェースケーブル5を介して接続されている。この様に、メインセンター9側で

はリモートコピー対象のデータが格納されるディスクサブシステム3について、1台のマスタディスクサブシステム3-1とその他のスレーブディスクサブシステム3-2、……3-nがインタフェースケーブル5で接続される構成をとる。

【0032】マスタディスクサブシステム3-1は、上位装置1がマスタディスクサブシステム3-1にデータの書き込み要求を発行すると、これに同期して当該データを自己のサブシステム内のデータバッファ12に書き込み、更に自己のサブシステム内のデータバッファ12にデータが書き込まれたこととは非同期に、遠隔地に存在するディスクサブシステム7-1に対し、データの書き込み指示を行うディスクサブシステムである。自己のサブシステム内のデータバッファ12に書き込まれた当該データは、同期或いは非同期にて磁気ディスクドライブ13に記録される。

【0033】非同期型の遠隔地にデータを書き込むリモートコピー方法として、メインセンター9のディスクサブシステム3は、自己のサブシステム内のボリュームが更新された順番に従い、自サブシステムが接続されているリモートセンタ10のディスクサブシステム7に更新データを送付し、リモートセンタ10のディスクサブシステム7は受け取った順番に従い、更新データを自己のサブシステム内のボリュームに反映するモードと、メインセンター9側が自己のサブシステム内のボリュームが更新された順番によらずディスクサブシステム3で最適にスケジューリングされた契機で送付対象のデータをまとめて送信し、リモートセンタ10のディスクサブシステム7は受け取った順番に関係なく更新データを自己のサブシステム内のボリュームに反映するモードを保有する。

【0034】スレーブディスクサブシステム3-2、……3-nは、上位装置1がディスクサブシステム3-2、……3-nにデータの書き込み要求を発行すると、これに同期して当該データを自己のサブシステム内のデータバッファ12に書き込み、さらにマスタディスクサブシステム3-1のリモートコピー制御情報格納部16の状態を参照し、リモートコピーの状態により、自己のサブシステム内のデータバッファ12にデータが書き込まれたこととは非同期に、遠隔地に存在するディスクサブシステム7-2、……7-nに対し、データの書き込み指示を行なうか、もしくは更新データの格納位置に関する情報を自サブシステム内のリモートコピー制御情報格納部16に保持するかを判断するディスクサブシステムである。ディスクサブシステム7-1、7-2、……7-nは、インターフェースケーブル4により接続されたディスクサブシステム3-1、3-2、……3-nから受け取ったデータを、自己のサブシステム内のデータバッファ12に格納する。

【0035】つまり、上位装置1から一台または複数台

のディスクサブシステム3-1、3-2、……3-nに対しデータの書き込み指示があった場合には、ディスクサブシステム3-1の状態次第で、リモートセンター10内の一一台または複数台のディスクサブシステム7-1、7-2、……7-nにも同じデータが格納されるシステム構成を示している。図1の矢印は、上位装置1から書き込み指示のあったデータの流れを示している。

【0036】なお、マスタディスクサブシステム3-1は、リモートコピーの状態を表す制御ビットをリモートコピー制御情報格納部16内に持っているので、システム運用者により予め設定された契機もしくは任意の時点でのシステム運用者による指示に基づき、この制御ビットの情報を変更することによってリモートコピーを一時停止状態にすることができる。リモートコピーが一時停止のときは、ディスクサブシステム3-1、3-2、……3-nは自己のディスクサブシステムのデータバッファ12へ更新データの格納を行うと共に、一時停止状態となった時点以後に受領した書き込み指示についての更新データの格納位置に関する情報をリモートコピー制御情報格納部16に保持しておき、ディスクサブシステム7-1、7-2、……7-nへの更新データの書き込み指示の発行は行わず保留する。

【0037】本願明では、このようにリモートコピーを一時停止させることにより、リモートセンター10側の全てのサブシステムにおいて、リモートコピーが一時停止させた時点のメインセンター9側のデータが存在する事となる。つまり、一時停止した時刻における、メインセンター9側のデータとリモートセンター10側のデータとの整合性を取ることが出来る。よって整合性を取るためにタイムスタンプをデータに付加させる必要が無くなり、時間情報が上位装置から付加されてこないオープン系のシステムであっても、上位装置の介在無しでリモートコピーを実現することができる。

【0038】また、ディスクサブシステム3-1、3-2、……3-nは、予めシステム運用者によりマスタディスクサブシステム3-1に設定された契機もしくは任意の時点でのシステム運用者による指示に基づき、上記一時停止状態を解除することができる。

【0039】一時停止状態が解除されると、ディスクサブシステム3-1、3-2、……3-nは、一時停止している間に更新のあったデータの書き込み指示をディスクサブシステム7-1、7-2、……7-nに対して発行するとともに、上位装置1からディスクサブシステム3-1、3-2、……3-nへデータの書き込み要求が発行された場合は、これに同期して当該データを自己のサブシステム内のデータバッファ12に書き込み、更に、自己のサブシステム内のデータバッファ12にデータが書き込まれたこととは非同期に、遠隔地に存在するディスクサブシステム7-1、7-2、……7-nに対し、データの書き込み指示を行う。

【0040】このような構成とする事により、メインセンター9内のリモートコピー対象のディスクサブシステム3のボリュームと、リモートセンター10内のディスクサブシステム7のボリュームにおいて、更新処理タイミングの遅れ等を無視できれば、同一のデータが保持される。さらに、マスタサブシステム3-1においてリモートコピーが一時停止状態となっている間は、マスタサブシステム3-1が一時停止状態となった時点のメインセンター9の各ディスクサブシステム3のデータの状態、つまり当該時点で一貫性が保証されたデータの状態が、リモートセンター10の各ディスクサブシステム7で保証・維持される。

【0041】なお、リモートコピーの一時停止や一時停止の解除は、リモートコピーのボリュームペア単位に設定できる。複数のボリュームペアを一つのボリュームグループに設定して、ボリュームグループ単位に状態を変化させることも可能である。そして、一時停止や一時停止解除を何れかのサブシステム3, 7や上位装置1, 8のコンソール、或いはこれらのシステムを管理する際に使用するモニターに表示することによって、ユーザーはリモートコピーが現在行われているか否か、またどのような単位でリモートコピーが行われているかを認識する事ができる。

【0042】このリモートコピーの一時停止及び一時停止解除の間隔は、一時停止される前の全データがリモートセンター10側にコピーされる前に、一時停止が解除され新しいデータがリモートセンター10側にコピーされてしまい、メインセンター9側とリモートセンター10側とで整合性がとれなくなる程の短時間で無い限り、ユーザーの任意に設定する事が可能である。

【0043】ここではリモートセンター10内においてリモートコピー一時停止状態の時点のサブシステム7のデータを保存するために行うコピーの時間をえた一例として、30分間メインセンター9からリモートセンター10へのリモートコピーを行い、次の30分間は一時停止、その後に一時停止を解除して再び30分間リモートコピーを行うというサイクルを挙げておく。もちろんリモートセンター10内のコピー時間が30分でない場合には、コピー時間に併せて一時停止をする時間を変えてよいし、コピー時間に囚われずに一時停止及び一時停止解除の間隔を設定してもよい。

【0044】また、マスタディスクサブシステム3-1内のリモートコピー制御情報格納部16に格納されたリモートコピー状態を表す制御ビットを、マスタディスクサブシステムがスレーブディスクサブシステム3-2、……3-nにインタフェースケーブル5を用いて予め配信しておくことにより、スレーブディスクサブシステムがマスタディスクサブシステムに制御ビットを問い合わせる必要を無くする構成としてもよい。その場合には、スレーブディスクサブシステムのリモートコピー制

御情報格納部16にマスタディスクサブシステムのリモートコピー制御情報格納部16と同様にスレーブサブシステム自身の状態管理も格納する。

【0045】上位装置8は、リモートセンター10においてディスクサブシステム7-1、7-2、……7-nとインターフェースケーブル6によって接続され、ディスクサブシステム7-1、7-2、……7-nに対し、参照及び更新を行う中央処理装置である。上位装置8は、メインセンター9の上位装置1が災害や故障等により本来の機能を果たせなくなった場合に、上位装置1の代替となって処理を行うことが出来る。このほか、ディスクサブシステム7-1、7-2、……7-nに格納されているデータを使用して、メインセンター9の上位装置1とは異なる処理を、上位装置1とは別個独立に実行することができるものである。

【0046】但し、上位装置8がディスクサブシステム7-1、7-2、……7-nに対し処理を行わない場合、上位装置1の代替機能を備えない場合には、上位装置8は不要である。逆に、上位装置8を備え、ディスクサブシステム7-1を他のディスクサブシステム7-2～7-nとインターフェースケーブル5で接続し、メインセンター9のマスタディスクサブシステム3-1と同様の構成とすることで図1のメインセンター9をリモートセンターに、リモートセンター9をメインセンターとして機能させる事も可能である。

【0047】本発明の実施の形態として、データの二重化方法と運用の概略を図2を用いて説明する。

【0048】二重化の対象となるデータが格納されたファイルやボリューム、ディスクサブシステム3は、事前に運用者が二重化つまりリモートコピーの必要に応じて選択する。そして、対象ファイルや対象ボリューム及びディスクサブシステム3と、選択したデータの複製を格納するファイルやボリューム及びディスクサブシステム7との関係や、二重化する際に更新順序の整合性を常時保持する必要があるか否かを、予め運用者が上位装置1或いはサービスプロセッサ15等からマスタディスクサブシステム3-1内のリモートコピー制御情報格納部16に対し設定しておく。通常は、データベースの更新履歴であるログファイル等のみに限って更新順序の整合性を常時保持する設定とする場合が多いが、本例ではファイルの種別に関わらず更新順序を常時保持する設定はしないこととする。

【0049】また、マスタサブシステム3-1については、リモートコピーを一時停止させる契機、並びに一時停止を解除する契機を設定する。契機の設定は、上位装置1から指示できるため、運用の自動化を支援する上位装置1のプログラムにより上位装置1からの指示契機を予めスケジューリングしておくことが可能である。

【0050】上記の選択、設定に際し、専用のコンソールやサービスプロセッサ15を接続又は装備できるディ

スクサブシステム3の場合には、上位装置1を利用せず、そのコンソールやサービスプロセッサ15を通じて設定できる。本例では、上位装置1を利用せず、ディスクサブシステム3内部で保有する時間値を利用して、予め運用者がマスタサブシステム3-1において、定期的にリモートコピーの一時停止、並びに一時停止の解除が実施されるように設定しておく。

【0051】図2のフローは専用のコンソールから選択・設定を行う場合を示している。なお、リモートコピーの一時停止や一時停止解除の設定は、リモートコピー対象のボリュームペア単位に設定する（ステップ1：図ではS1の様に示す。以下同じ）。通常は、リモートコピー対象のボリュームペアすべてを一つのボリュームグループとして定義し、ボリュームグループ内のボリュームはすべて同一のステータスとなるように設定する。

【0052】本例では、ディスクサブシステム3のボリューム全てをリモートコピー対象とする。従って、以下では、リモートコピーの状態を、ボリュームペアやボリュームグループ単位ではなく、ディスクサブシステム単位として記述する。なお、本例では詳細は記述しないが、データベースとログファイルとはボリュームグループを分けて設定し、ログファイルを格納するボリュームについては、リモートコピーの一時停止契機や一時停止解除契機を設定しないといった定義も可能である。

【0053】リモートコピー対象のファイルやボリュームの設定方法としては、ボリュームやディスクサブシステムを意味する具体的なアドレスを指定する方法や、ディスクサブシステム内の制御プログラムによって、アドレスの任意の範囲から選択する方法を取ることもできる。初期設定として、バス設定やペア設定、並びに一時停止契機や一時停止解除契機の設定を行う例を示してある。

【0054】上位装置1から、ディスクサブシステム3-1に対し、データの書き込み要求（以下、ライトコマンド）が発行される（ステップ2）と、ディスクサブシステム3-1はライトコマンドにもとづき自己のディスクサブシステム内へデータ格納処理を実行し（ステップ3）、自己のサブシステム内へのデータ書き込み（格納）処理が完了した後に、上位装置1に対しライトコマンドに対する処理の完了報告を行う（ステップ4）。

【0055】また、上位装置1から、ディスクサブシステム3-2、……、3-nに対し、ライトコマンドが発行されると（ステップ2）、ディスクサブシステム3-2、……、3-nはライトコマンドに基づき自己のディスクサブシステム内へデータ格納処理を実行する（ステップ5）。ここで、ライトコマンドとは、データを書き込むための指示と書き込みデータそのものを転送するコマンドであり、どのディスクサブシステムに要求を出すのかはユーザが予め上位装置1に対して設定しておく（ステップ1）。

【0056】ライトコマンドを受領した際、ディスクサブシステム3-1は、自サブシステムのリモートコピーの状態を表すリモートコピー制御情報格納部16内の制御ビットを参照し、自サブシステムのリモートコピー状態を判別する（ステップ6：図3）。自サブシステムがリモートコピー一時停止状態であった場合、ディスクサブシステム3-1は自サブシステムに接続されているリモートセンター10のディスクサブシステム7-1へは更新データの送信は行わず、更新が実施されたデータの格納位置に関する情報を自サブシステム内に保持しておく（ステップ7）。自サブシステムがリモートコピー一時停止状態になれば、自己のサブシステムの処理能力に基づいて決定された契機で、ライトコマンドをディスクサブシステム7-1に対し発行する（ステップ8）。

【0057】なお、リモートコピー一時停止状態の間に更新が行われたデータの格納位置情報を保持している場合は、当該位置のデータもリモートセンター10のディスクサブシステム7-1への送付対象と判断し、当該データを書き込む為のライトコマンドを発行し、ライトコマンドに対する処理が完了した後に更新位置情報を消去する。

【0058】ライトコマンドを受領した際、ディスクサブシステム3-2、……3-nは、ディスクサブシステム3-1に対しインタフェースケーブル5を経由してディスクサブシステム3-1の状態を問い合わせるコマンドを発行し、ディスクサブシステム3-1のリモートコピー状態を表す制御ビットを取得・参照する（ステップ9）ことにより、ディスクサブシステム3-1がリモートコピー一時停止状態にあるか否かを確認する（ステップ10）。

【0059】ディスクサブシステム3-2、……3-nはディスクサブシステム3-1がリモートコピー一時停止状態であった場合、更新が実施されたデータの格納位置に関する情報を自サブシステム内に保持（ステップ12）しており、上位装置1に対しライトコマンドに対する処理の完了を報告する（ステップ13）。

【0060】また、ディスクサブシステム3-1がリモートコピー一時停止状態でない場合は、ディスクサブシステム3-2、……3-nは、上位装置1に対しライトコマンドに対する処理の完了を報告し（ステップ14）、自己のサブシステムの処理能力に基づいて決定された契機で、ライトコマンドをディスクサブシステム7-2、……7-nに対し発行する。なお、リモートコピー一時停止状態の間に更新が行われたデータの格納位置情報を保持している場合は、当該位置のデータもリモートセンタのディスクサブシステム7-2、……7-nへの送付対象と判断し、当該データを書き込む為のライトコマンドを発行し（ステップ15）、ライトコマンドに対する処理が完了した後に更新位置情報を消去する。

【0061】つまり、ディスクサブシステム3-1がリモートコピー一時停止状態に有れば、ディスクサブシステム3-1に接続されているメインセンタ9の他のディスクサブシステムは上位装置1からライトコマンドが発行される事に起因しすべてリモートコピー一時停止状態となる。また、ディスクサブシステム3-1がリモートコピー一時停止状態でなければ、ディスクサブシステム3-1に接続されているメインセンタ9の他のディスクサブシステムは上位装置1からライトコマンドが発行される事に起因しリモートコピーが実施される。

【0062】尚、ディスクサブシステム3-2、……3-nがディスクサブシステム3-1に問い合わせる（ステップ9）代わりに、ディスクサブシステム3-1が自らのリモートコピー状態が変化していた場合に、ディスクサブシステム3-2、……3-nがディスクサブシステム3-1にその変更を伝える（ステップ9'：図示せず）ような設定、或いは先にも述べたように、ディスクサブシステム3-1が自らのリモートコピー状態が変化した際にディスクサブシステム3-1がディスクサブシステム3-2、……3-nにその変更を伝える設定としてもよい。

【0063】この様な設定とした場合には、ディスクサブシステム3-2、……3-nにおいても、ディスクサブシステム3-1と同様に自らのリモートコピーの状態を保存しておく必要があるため、ステップ10においてディスクサブシステム3-1がリモートコピー一時停止状態であった場合、自ディスクサブシステムのリモートコピー状態を一時停止状態に変更する（ステップ11：図示せず）。また、ステップ10においてディスクサブシステム3-1がリモートコピー一時停止状態でない場合、自ディスクサブシステムのリモートコピー状態を一時停止状態解除に変更する（ステップ11'：図示せず）。

【0064】尚、ステップ9のようにディスクサブシステム3-2、……3-nがディスクサブシステム3-1に問い合わせる設定であっても、自ディスクサブシステムのリモートコピー状態を表示するなどしたい場合は、自ディスクサブシステム内にリモートコピー状態を保存し、ステップ11及びステップ11'を設けても良い。

【0065】ディスクサブシステム7-1、7-2、……7-nは、ディスクサブシステム3-1、3-2、……3-nから発行されたライトコマンドを受領していることを確認すると、ライトコマンドに対する処理、即ち、自己のサブシステム内のデータバッファ12へのデータ格納処理を行う（ステップ16）。

【0066】ディスクサブシステム7-1、7-2、……7-nは、ライトコマンドに対する処理、即ち、自己のサブシステム内のデータバッファ12へのデータ格納処理が完了すると、ディスクサブシステム3-1、

3-2、……3-nに対し、ライトコマンドに対する処理完了報告を行う（ステップ17）。

【0067】なお、一時停止状態が解除されると、ディスクサブシステム3-1、3-2、……3-nは、自サブシステムのリモートコピーの状態が一時停止状態となつた時点以後に更新のあったデータの格納位置情報をもとに、当該位置のデータの書き込み指示をディスクサブシステム7-1、7-2、……7-nに対して発行するとともに、上位装置1からディスクサブシステム3-1、3-2、……3-nへデータの書き込み要求が発行された場合は、これに同期して当該データを自己のサブシステム内のデータバッファ12に書き込み、更に、自己のサブシステム内のデータバッファ12にデータが書き込まれたこととは非同期に、遠隔地に存在するディスクサブシステム7-1、7-2、……7-nに対し、データの書き込み指示を行う動作となる。

【0068】本発明により、上位装置1から書き込まれたデータは、ディスクサブシステム3-1、3-2、……3-nに格納されるだけでなく、ディスクサブシステム7-1、7-2、7-nにも複写され格納される。また、ディスクサブシステム3-1がリモートコピー一時停止状態となつた時点のディスクサブシステム3-1、3-2、……3-nのデータの状態が、リモートセンタ10側のディスクサブシステム7-1、7-2、……7-nで生成される。メインセンター9が被災した場合は、ディスクサブシステム7-1、7-2、……7-nのデータを利用して、ジョブを再実行する等の回復作業を行い、業務を再開する。

【0069】なお、リモートセンター10側では、リモートコピーが一時停止状態の間に、ディスクサブシステムのボリューム複写機能等を用いて、ディスクサブシステム7-1、7-2、……7-nのデータの複製を作成し保存しておく。複製を作成しておくことにより、リモートコピー実行中にメインセンター9が被災し、ディスクサブシステム7-1、7-2、……7-nにデータが書き込み中であったためにデータの一貫性が損なわれた場合にも、保存しておいた当該データの複製をもと

にジョブを再実行する等の回復作業が行える。

【0070】これらはすべてディスクサブシステムの機能のみで実現され、上位装置の処理能力に対し負担となるない。

【0071】

【発明の効果】上位装置への新規ソフトウェアの導入を必要とせずサブシステム側の機能変更のみで、ユーザが期待する範囲での更新データの一貫性を保証でき、導入が容易なリモートコピーシステムを実現できる。

【0072】また、メインセンターとリモートセンターとの間のデータ送受信の際は、データはまとめられた単位で処理されるため、データ転送効率を向上できるだけでなく、更新順序の整合性を保持するための制御を行うことに起因するサブシステムの制御オーバヘッドも削減できるので、高性能で導入の容易な災害バックアップシステムを提供することができる。

【図面の簡単な説明】

【図1】本発明の一実施の形態におけるリモートコピーシステムの全体構成を示す図である。

【図2】リモートコピーシステムの処理を示すフローチャートである。

【図3】図2の続きのリモートコピーシステムの処理を示すフローチャートである。

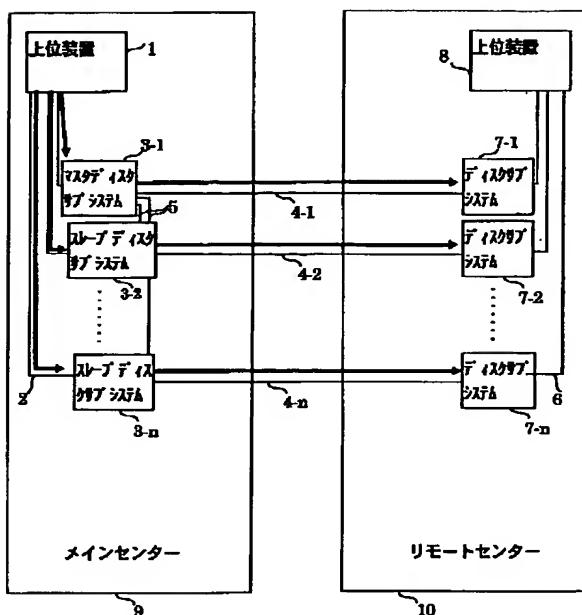
【図4】メインセンターのメインディスクサブシステム及びスレーブディスクサブシステムの構成を示す図である。

【符号の説明】

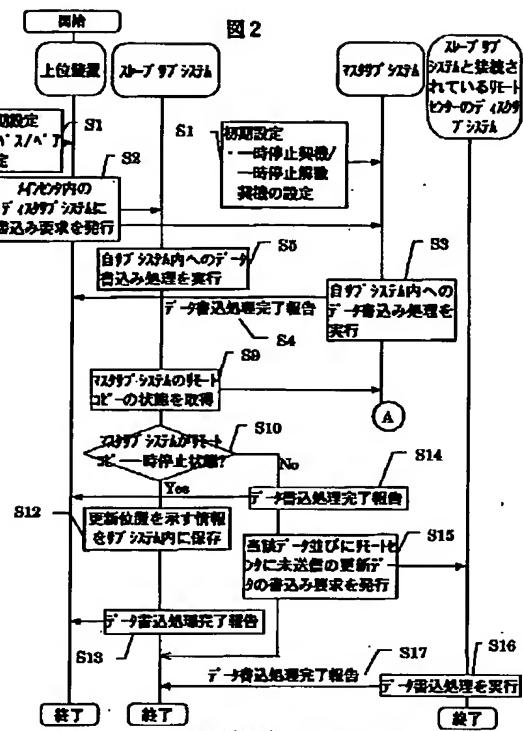
1…上位装置、2…インタフェースケーブル、3…ディスクサブシステム、4…インタフェースケーブル、5…インタフェースケーブル、6…インタフェースケーブル、7…ディスクサブシステム、8…上位装置、9…メインセンター、10…リモートセンター、11…インタフェース制御部、12…データバッファ、13…磁気ディスクドライブ、14…マイクロプロセッサ、15…サービスプロセッサパネル、16…リモートコピー制御情報格納部、17…ディスクアレイサブシステム制御部。

【図1】

図1

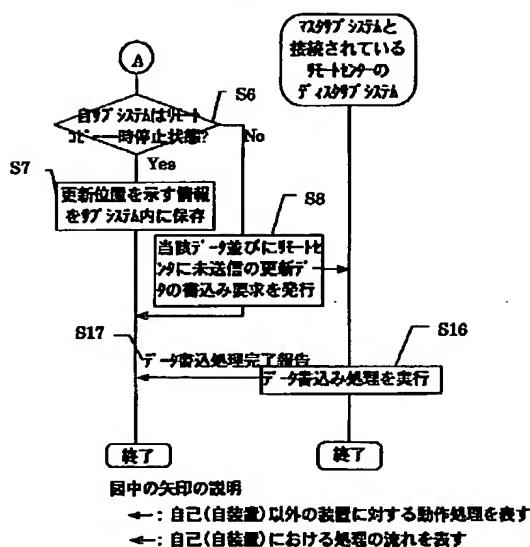


【図2】



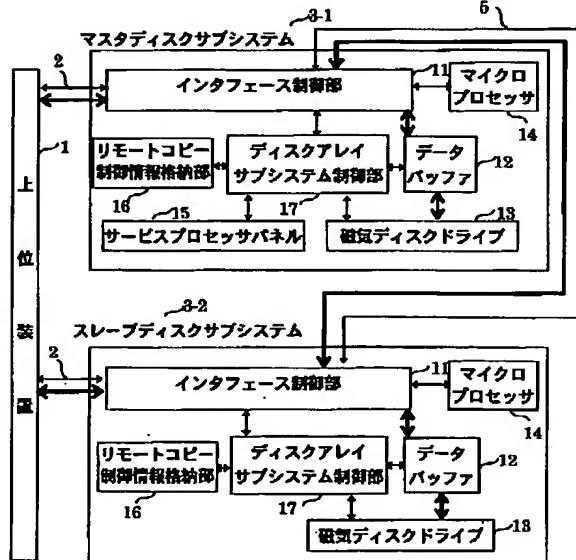
【図3】

図3



【図4】

図4



↔ : 制御情報の流れ
↔↔ : データの流れ

フロントページの続き

(72)発明者 野沢 正史

神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内

(72)発明者 中野 傑夫

神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内
F ターム(参考) 5B065 BA01 CA12 CE22 EA12 EA33
5B082 DE07